## Final Project
TA email: `mltech_ta@csie.ntu.edu.tw`

RELEASE DATE: 03/29/2020

COMPETITION END DATE: **06/16/2020 23:59:59 ONLINE**

REPORT DUE DATE: **06/26/2020 NOON ONLINE**

*Unless granted by the instructor in advance, no late submissions will be allowed.*

*Any form of cheating, lying, or plagiarism will not be tolerated. Students can get zero scores and/or fail the class and/or be kicked out of school and/or receive other punishments for those kinds of misconducts.*

*You should write your solutions in English or Traditional Chinese with the common math notations introduced in class or in the problems. We do not accept solutions written in any other languages.*

## Introduction

In this final project, you are going to be part of an exciting machine learning competition. The competition is here and we hope that you like mangoes!

<div align="center">

`https://aidea-web.tw/aicup_mango`
`https://aidea-web.tw/topic/72f6ea6a-9300-445a-bedc-9e9f27d91b1c`

</div>

We ask you to attend the preliminary round of the grade competition. It is totally up to you whether to attend other rounds/competitions to win the prize. You need to follow the competition rules and fight for the leading positions on the score board. To pass this course, you need to submit a comprehensive report that describes not only the recommended approaches, but also the reasoning behind your recommendations. Well, let's get started!

## Survey Report

You are asked to study at least THREE machine learning approaches. Then, you should make a comparison of those approaches according to some different perspectives, such as efficiency, scalability, popularity, and interpretability. In addition, you need to recommend THE BEST ONE of those approaches as your final recommendation and provide the "cons and pros" of the choice.

The survey report should be no more than SIX A4 pages with readable font sizes. The most important criterion for evaluating your report is replicability. Thus, in addition to the outlines above, you should also describe how you pre-process your data, such as the features you build; introduce the approaches you tried and provide specific references, especially for those approaches that we didn't cover in class; list your experimental settings and the parameters you used (or chose) clearly. Other criteria for evaluating your survey report would include, but are not limited to, clarity, strength of your reasoning, "correctness" in using machine learning techniques, the work loads of team members, and properness of citations.

> Our sincere suggestion: *Think of your TAs as your boss who wants to be convinced by your report.*

For grading purposes, a minor but required part in your survey report for a two- or three-people team (see the rules below) is how you balance your work loads.

## Competition

You should simply go to the official site to form the team and to submit. The competition can be registered before 05/15/2020 and ends at 23:59:59 on 06/16/2020. We encourage you to register as early as possible.

# Misc Rules

**Report**: Please upload one report per team electronically on Gradescope. You do not need to submit a hard-copy. The report is due at noon on 06/26/2020.

**Teams**: By default, you are asked to work as a team of size THREE. A one-person or two-people team is allowed only if you are willing to be as good as a three-people team. It is expected that all team members share balanced work loads. Any form of unfairness, such as the intention to cover other members' work, is considered a violation of the honesty policy and will cause some or all members to receive zero or negative score.

**Algorithms**: You can use any algorithms, regardless of whether they were taught in class.

**Packages**: You can use any software packages for the purpose of experiments, but please provide proper references in your report for replicability.

**Source Code**: You do not need to upload your source code for the final project. Nevertheless, please keep your source code until 08/01/2020 for the graders' possible inspections.

**Grade**: The final project is worth 400 points. That is, it is equivalent to two usual homework sets. At least 360 of them would be reserved for the report. The other 40 may depend on some minor criteria such as your competition results, your discussions on the boards, your work loads, etc..

**Collaboration**: The general collaboration policy applies. In addition to the competitions, we still encourage collaborations and discussions between different teams.

**Data Usage**: You can use only the data sets provided in the competition and follow the rules of the organizers, and you should use the data sets properly. Getting other forms of the data sets is strictly prohibited and is considered a serious violation of the honesty policy. Using any tricks to query the labels of the test set is also strictly prohibited.